National Aeronautics and Space Administration

PBS Queues and Scheduling on Pleiades

June 11, 2013

NASA Advanced Supercomputing Division

Outline

- Overview of Pleiades Queues
  - ~40 queues , but only a few queues matter to each user
- Reading the Fine Prints about a Queue
  - Access control, defaults, limits, and status
- Queues with Restrictive Access
  - ***vlong***, ***wide***, special, and reservation queues
- Resources and Jobs Monitoring Tools
  - *qstat* options, *node_stats.sh*, and *qs*
- How PBS Job Scheduling Works
  - It's complicated
  - General sorting
  - Typical scheduling cycle

Question? Use the Webex chat facility to ask the Host

Overview of Pleiades Queues

- *% qstat –Q* shows basic info for each of the ~40 queues

```
Queue           Ncpus/       Time/        State counts
name              max/def     max/def     T/___Q/_H/W/__R/E/B pr  Info
=============== =====/=== ======/====== ================== === ========
somd_spl           --/  8 240:00/ 01:00 0/___0/_0/0/__0/0/0  25 disabled
smd1               --/  8 240:00/    -- 0/__48/_0/0/__0/0/0  16
.....
```

Ncpus/max: not set for almost all queues

Time/max : 2 hours to 16 days

pr: priority (ranging from -10 to 149 for existing queues)

- Most queues can be ignored
  - Staff testing queues: ***alphatst, ded_time, diags, dpr, idle, resize_test, testing***
  - Routing queue: ***route*** queue routes Pleiades jobs to ***normal*** or ***long***
  - Queues for Pleiades GPU: ***gpu***, ***gpu_long_free***
  - Queues for Endeavour: ***e_normal, e_long, e_debug, e_low***

Question? Use the Webex chat facility to ask the Host

Overview of Pleiades Queues (cont'd)

- Queues that all Pleiades users can submit to
  - ***normal, long, debug, devel (***and ***low)***
- Queues that all Lou2 users can submit to (from lfe1/2 or pfe's)
  - ***ldan***
- Queue for very long (5-16 days) jobs
  - ***vlong***
- Queue for wide (> ¼ mission shares or model type) jobs
  - ***wide***
- Special queues for some groups or users
  - ***armd_spl, heomd_spl, smd_ops, iss, kepler***, etc
- Reservation queues
  - ***R1155585*** (for example)

Question? Use the Webex chat facility to ask the Host

Reading the Fine Prints about a Queue

- *% qstat –fQ queue_name*

  displays various attributes set for the queue
- Meaning of each attribute can be found in

  pfe: /PBS/doc/PBSProRefGuide11.2.pdf
  - queue_type: Execution or Route
  - priority: larger number -> higher priority
  - access control: for groups/users; if enabled, a list of groups/users is shown;

    if both enabled, both have to be satisfied
  - defaults for a job: model, ncpus, walltime
  - limits for a job: min or max on ncpus, walltime, nodect, etc.
  - status: # of jobs in various states; amount of various resources (mem, ncpus, nodect, nobackup, etc) assigned to running jobs; is the queue enabled/started?

    enabled: queue will accept jobs
    started: jobs may be scheduled to run

```
pfe20% qstat -fQ normal
Queue: normal
    queue_type = Execution
    Priority = 0
    total_jobs = 499
    state_count = ..Queued:421 Running:43...

    acl_group_enable = True
    acl_groups = a0709,a0801,a0805,.....
                 .........
    resources_default.ncpus = 8
    resources_default.walltime = 01:00:00
    default_chunk.model = har

    resources_min.ncpus = 1
    resources_max.walltime = 08:00:00

    resources_assigned.mpiprocs = 1961
    resources_assigned.ncpus = 4186
    resources_assigned.nobackupp1 = 21
    resources_assigned.nobackupp2 = 19
    resources_assigned.nobackupp4 = 0
    resources_assigned.nobackupp6 = 3
    resources_assigned.nodect = 497

    enabled = True
    started = True
```

Queues for All Pleiades Users

Queue name	Ncpus/ max/def	Time/ max/def	pr
low	--/ 8	04:00/ 00:30	-10
normal	--/ 8	08:00/ 01:00	0
long	--/ 8	120:00/ 01:00	0
debug	--/ 8	02:00/ 00:30	15
devel	--/ 1	02:00/ --	149

- Purpose
  - ***low, normal, long*** – for production runs
  - ***debug, devel*** – debug or development (pr > 0)
- Access
  - acl_group_enable=True for all these queues
  - acl_groups= list of all gids with > 0 allocation
  - New gids are added by a script once a day

- Other Defaults
  - default_chunk.model: **wes** (westmere) for ***devel***, har (harpertown) for the other four
  - All four models (har,neh,wes,san) are available for these queues

    Specify **:model=xxx (xxx=har,neh,wes,san)** if you want to use non-default model
- Other Limits

low: max_run = [u:PBS_GENERIC=100] — max # of running jobs per user is 100

long: resources_max.nodect = 1024 — max # of nodes per job Is 1024

max_run_res.nodect = [o:PBS_ALL=11250] — max of a total of 11,250 nodes for all running jobs

debug: resources_max.nodect = 128 — max # of nodes per job is 128

max_run=[u:PBS_GENERIC=2] — max # of running jobs per user is 2

max_run_res.nodect = [u:PBS_GENERIC=128] — max # of nodes per user is 128 (total of 2 running jobs)

devel: max_queued = [u:PBS_GENERIC=1] — max # of job queued or running per user is 1

resources_max.nodect = 512 — max # of nodes per job is 512

Question? Use the Webex chat facility to ask the Host

Queues for All Pleiades Users (cont'd) - *devel* queue

- **Purpose: code, script, parameters, datacase development**
- Provides fastest access to resources (priority = 149) for development work
  - wait time < 1 min for >70% of jobs
- Resources are made available dynamically based on demand
  - during prime-time (4AM-7PM PDT), PBS attempts to maintain more free nodes for each model to accommodate more devel work
  - through a node shuffling mechanism (using PBS jobs submitted by staff **gamatthe** with jobname **node_shuffle**)
- Devel jobs can start even if mission shares have been exceeded
  - but resources used by devel jobs are counted as usage against mission shares
  - will affect other jobs of the same mission
- **DO NOT ABUSE**
  - NAS is developing tools to detect possible abuse
- Additional Info

  http://www.nas.nasa.gov/hecc/support/kb/Pleiades-devel-Queue_290.html

Question? Use the Webex chat facility to ask the Host

Queue for All Lou2 Users - *ldan* Queue

- LDAN: Lou Data Analysis Nodes (10 Sandy Bridge nodes total, underused)
- Access
  - pfe, bridge, lfe1, lfe2, lou2% qsub –q ***ldan*** job_script
  - Lou1 users have to wait till Lou1/2 are merged
- Default
  - default_chunk.model = **ldan**
    Job will never start if you specify **:model=har,neh,wes,san**
- Purpose
  - data processing that needs more memory: ldan1-5: 96 GB, ldan6-10: 256 GB
  - mainly for post-processing data under lou2: /u/username
  - also allows pre- or post-process data under pfe:/nobackup/username
    Be aware that $HOME on ldan nodes is Lou2's $HOME, not Pleiades $HOME.
    Your pfe: $HOME/.cshrc is not sourced on ldan nodes.
- Limits
  - resources_max.nodect = 1 — max # of nodes per job is 1
  - resources_max.walltime = 72:00:00 — max walltime per job is 3 days
  - max_run = [u:PBS_GENERIC=2] — max # of running jobs per user is 2
- Additional Info
  http://www.nas.nasa.gov/hecc/support/kb/Lou-Data-Analysis-Nodes_413.html

Question? Use the Webex chat facility to ask the Host

Queues with Restrictive Access - *vlong* Queue

- Purpose
  - for jobs that need to run for more than 5 days
- Access
  - acl_group_enable = True
  - acl_user_enable = True
    Access granted by NAS User Services Manager upon request
- Default
  - default_chunk.model = **har**
    specify **:model=neh,wes,san** if you want other model
- Limits
  - resources_max.nodect = 1024 — max # of nodes per job is 1024
  - resources_max.walltime = 384:00:00 — max walltime per job is 16 days
  - max_run_res.nodect = [o:PBS_ALL=11250] — max of a total of 11,250 nodes for all running jobs
- Checkpoint/restart
  - implement checkpoint/restart in your code (for vlong and all other queues)

Question? Use the Webex chat facility to ask the Host

Queues with Restrictive Access - *wide* Queue

- Purpose
  - To decrease the wait time of a 'wide' job by giving it a higher priority (pr=101)
  - 'wide' # of total ncpus requested > ¼ of your mission shares (in NCPUS)
    or # of nodes requested > ¼ of total nodes of a model type

Sample output taken from
pfe20% *qstat –W shares=-*

Group	Share %	Use%	Share (in cores)	1/4
ARMD	25	18	49,376	12,344
HEOMD	24	18	46,858	11,714
SMD	47	44	92,556	23,139

Share %: subjects to change, as directed by HECC manager
Use%: changes dynamically
Share: depends on share % and total resources that are not down or offline

# of Nodes	Total	¼
Har	4,096	1,024
Neh	1,280	320
Wes	4,672	1,168
San	1,872	468

Assuming no nodes are down or offlined

```
pfe20% qstat –a -W o=+model,mission,pri wide
130000.pbspl1 username wide jobname  4096  512  24:00 Q 3d+14:51 -- 512:neh SMD 101
```

#PBS –lselect=512:ncpus=8:model=neh ← request 512 Nehalem nodes and a total of 4096 cores

Question? Use the Webex chat facility to ask the Host

Queues with Restrictive Access - *wide* Queue (cont'd)

- Access
  - enabled = False

  PBS will not accept your job if you submit to it.
  Instead, a wide job is qmoved to it by staff

- Limits
  - max_run = [u:PBS_GENERIC=1]

  max # of jobs running (per user) is 1

  Only 1 job (total on the system) is allowed (by staff) to run in the ***wide*** queue.
  If more than 1 jobs are qualified, preference is given to
  * the one that has been waiting for longer time in other queues
  * the one whose owner does not have other jobs running

- Default
  - default_chunk.model = **har**

    specify **:model=neh,wes,san** if you want other model

Question? Use the Webex chat facility to ask the Host

Queues with Restrictive Access
- *various special Queues*

- Purpose
  - To allow time-critical projects to be processed with higher priority (pr varies)
    Other projects in the same mission may experience longer wait time

- Access
  - acl_group_enable or acl_users_enable = True

  Creation of and access to a special queue is authorized by the appropriate Mission or Program POC (the one who approves your allocation in e-Books).

  Mission Directorates are reluctant to authorize use of special queue lately.

Question? Use the Webex chat facility to ask the Host

Queues with Restrictive Access - *Advanced Reservation Queues*

- Purpose
  - To run extra-large job (>15,000 cores or as determined by HECC manager) that will be almost impossible to start without a reservation
- Access
  - Authorized by NAS HECC manager
  - Created by NAS staff using *pbs_rsub* command for the authorized user

    pfe20% *pbs_rsub –G s1020 –R 1130 –D 2:00:00 –lselect=1000:ncpus=12*

    This command attempts to reserve 1000 nodes for GID s1020 with a start time at 11:30 AM and a duration of 2hrs.

    If successful, pbs_rsub returns a queue name R****** (for example, R1155585)
  - Accessed by user: *qsub –q R****** job_script*
- SBU Charging
  - Charging applies to the whole reservation period, even if not used.

Question? Use the Webex chat facility to ask the Host

Demand for NAS Resources

- NAS supports more than 400 projects, 700 users
- PBS server status
  - Common to see hundreds of jobs waiting, sometimes more than the jobs running

```
pfe20% qstat –B
Server                 State count        Default
name                T/__Q/_H/W/__R/E/B queue   Info
=================== ================== ======= ==========
pbspl1.nas.nasa.gov 0/699/89/0/500/0/0 route   Scheduling
```

- Suggestions for increasing your job's chances of starting sooner
  - Be flexible with model type
  - Use fewer nodes
  - Request < 4 hrs to allow borrowing from other missions
  - Use the shrink-to-fit feature when you can

    #PBS -l min_walltime=2:00:00,max_walltime=8:00:00

    http://www.nas.nasa.gov/hecc/support/kb/Running-Jobs-Before-Dedicated-Time_306.html

Question? Use the Webex chat facility to ask the Host

Resources Monitoring Tools

- pfe20% *qstat –au foo*
  - Shows raw info for different model types in segments
  - Shows numbers of nodes (=hosts), CPUs and info on whether it is unused (blank), in-used, down, offline, etc.

```
pfe20.ychang1 83> qstat -au foo
pbspl1.nas.nasa.gov:                  Mon Jun  3 11:17:13 2013
 Server reports 1288 jobs total (T:0 Q:699 H:89 W:0 R:500 E:0 B:0)

  Host          CPUs Tasks Jobs Info
  ---------- ----- ----- ---- ------------------------
    176 hosts  1408     0   -- har
     48 hosts   384     0   -- har bigmem
     16 hosts   128   128   -- har bigmem in-use
  r32i2n11        8     0    0 har bigmem q=diags
      9 hosts    16     0   -- har down
   3770 hosts 30160 25057   -- har in-use
     11 hosts     0     0   -- har offline
     64 hosts   512     0   -- har q=devel
  r31i3n0         8     0    0 har q=smd_ops offline
      9 hosts   144     0   -- ldan q=ldan
  ldan10          0     0    0 ldan q=ldan offline
     25 hosts   400     0   -- neh
     13 hosts     0     0   -- neh down
   1128 hosts 18048  8959   -- neh in-use
      4 hosts    16     0   -- neh offline
     32 hosts   512     0   -- neh q=devel
     17 hosts   272   136   -- neh q=devel in-use
  r162i2n7       16     0    0 neh q=diags
     60 hosts   960   960   -- neh q=smd_ops in-use
  r311i4n7       16     0    0 san down
  r327i6n8       16     0    0 san down
   1622 hosts 25952 21849   -- san in-use
```

Harpertown 4096 nodes

Nehalem 1280 nodes

Question? Use the Webex chat facility to ask the Host

Resources Monitoring Tools

- node_stats.sh
  - Shows numbers of nodes (not including nodes that are down or offlined)
  - Includes numbers of nodes (for each model) queued jobs are waiting for
  - Does not include Mission Shares usage info

```
pfe20% /u/scicon/tools/bin/node_stats.sh

 Pleiades Nodes Total: 11876
 Pleiades Nodes Used : 10558
 Pleiades Nodes Free : 1318

 Harpertown               Total:  4075, Used:  3786, Free:   289
 Nehalem                  Total:  1263, Used:  1205, Free:    58
 Westmere                 Total:  4612, Used:  3709, Free:   903
 SandyBridge              Total:  1866, Used:  1803, Free:    63
 GPU nodes                Total:    60, Used:    55, Free:     5
 Devel har                Total:    64, Used:     0, Free:    64
 Devel neh                Total:    49, Used:    17, Free:    32
 Devel wes                Total:   714, Used:   200, Free:   514
 Devel san                Total:   206, Used:   168, Free:    38
 Devel bigmem(har)        Total:     0, Used:     0, Free:     0
 Devel bigmem(wes-48g)    Total:     0, Used:     0, Free:     0
 Devel bigmem(wes-96g)    Total:     0, Used:     0, Free:     0

Currently queued jobs are requesting:  3584 Harpertown,  1210
Nehalem,  8926 Westmere,  4552 SandyBridge nodes
```

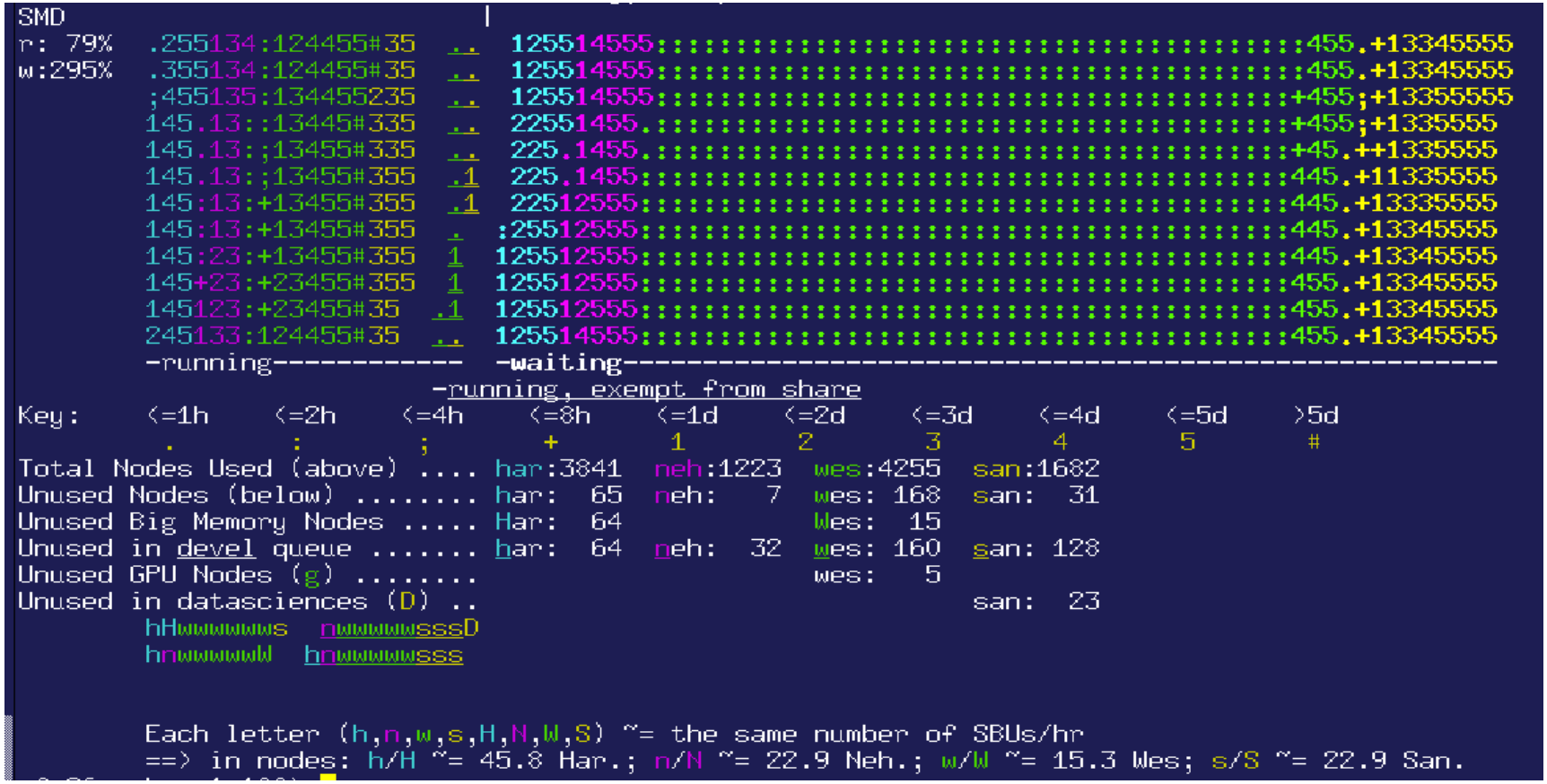
Include those in the Devel queue

Devel queue

Question? Use the Webex chat facility to ask the Host

Resources Monitoring Tools

- /u/scicon/tools/bin/qs
  - provides a graphical view of resources used and waited for in each mission
  - shows processor model types in different colors
  - shows how long the resources are tied up (in symbols and numbers of days) for running jobs or requested for queued jobs
  - provides a summary of nodes used and unused for each model



Question? Use the Webex chat facility to ask the Host

Jobs Monitoring Tools

- *qstat –nu username*: shows jobs of username

  where –n shows the exec_host string (good to know for running jobs)
- *qstat –s jobid* : gives brief message why job is not started
- *qstat –i* : shows jobs that are not running, in rough order PBS will process them
- *qstat –i –W o=+mission | grep ARMD*: shows ARMD that are not running
- *qstat –i –W o=+model | grep san*: shows jobs waiting for Sandy Bridge nodes

Question? Use the Webex chat facility to ask the Host

Scheduling – Job sorting

- The scheduler uses a multi-level sort to get its first approximation of sort order:
  - Mission share
  - Job priority
  - Queue priority
  - Job node count
- Then there are exceptions, special cases, and nuances…

- Mission share
  - Primary sort key
  - Mission that is using the smallest percentage of its share is first
  - Highly dynamic at times
  - Accomplishes NAS goal of meeting agreed-upon targets for NASA Mission Directorates

Question? Use the Webex chat facility to ask the Host

Scheduling – Job sorting (cont'd)

- Job priority
  - Typical job has priority 0, unless:
    - +10 boost when a user has no running jobs
    - +1 boost every 12 hours waited, capped at +20
  - Then there are exceptions, special cases, and nuances…

- Queue priority
  - Typical queue has priority 0
  - Special case queues have varying priorities greater than 0

- Job node count
  - Jobs with a higher node count sort higher
  - Assists with making sure larger jobs don't starve for resources due to smaller jobs packing most/all resources

Question? Use the Webex chat facility to ask the Host

20

Other Scheduling Features

- Top jobs
  - Certain jobs that cannot be run immediately are given a virtual reservation by the scheduler, setting aside particular nodes the scheduler predicts will be used to run the job in the future
  - Top jobs entail a high scheduling cost
  - Each mission is given at least 2 top jobs
  - Scheduler typically does 10-15 top jobs each cycle
- Backfill
  - Whenever possible the scheduler runs jobs that will finish before an upcoming reservation (virtual or real) or dedicated time
- Fairshare
  - Job sort order can result in “ties”
  - When two jobs have the same sort order, generally the job belonging to the user who has used less computing time sorts higher
  - Computing time usage is tracked and decays by cutting the value in half every 24 hours

Question? Use the Webex chat facility to ask the Host

Typical Scheduling Cycle

- The scheduler starts by asking the PBS server for the current state of the world
  - This state is assumed for the duration of the cycle
- Devel queue jobs are considered
  - Devel queue has a partitioned set of nodes
- ”Node shuffle” jobs are considered
  - Mechanism to grow the devel queue
  - Hence the limited scope of the devel queue
- Special jobs are considered (e.g. wide queue)
- Top jobs are identified and considered
- Remaining jobs are considered

- Note: sort order applies to each of the steps above

Question? Use the Webex chat facility to ask the Host